INDUSTRIAL SPILLOVERS IN DEVELOPING COUNTRIES: PLANT-LEVEL EVIDENCE FROM CHILE, MEXICO AND MOROCCO

By

C.J. Krizan*
The Center for Economic Studies
U.S. Bureau of the Census
Washington, DC 20233

CES 98-2   January 1998

DC 20233-6101, (301-457-1882) or INTERNET address
snguyen@info.census.gov.

Abstract

Recent trade and growth models have underscored the potential importance of external economies of scale. However, many of the most frequently modeled externalities have either not been measured or have been estimated with data too aggregate to be informative. In this paper, plant-level longitudinal data from Chile, Mexico and Morocco allow me to provide some of the first micro evidence on several types of external economies from plant-level production functions. The results indicate that in many industries own-industry output contributes positively to plant-level productivity. However, the effects of geographic concentration are mixed. Cross-country concentration, as measured by a geographic GINI index, often decreases productivity but within-province, same industry activity enhances it.

KEYWORDS: External Economies, Spillovers, Developing Countries, Microdata

*****This paper is a revised version of my Ph.D. dissertation at Georgetown University. I'd like to acknowledge the substantial assistance generously given me by Jim Tybout and Dan Westbrook. All mistakes are the responsibility of the author alone.

I. Introduction:

In many contexts the productivity of plants is positively correlated with the volume of economic activity—industry-wide, region-wide, economy-wide, or world-wide. When individual plants are not compensated for their contribution to this productivity effect, they do not factor it into their decision-making and external economies are said to be present. Recent analytical models have demonstrated that this phenomenon can critically influence the direction and welfare effects of international trade or encourage self-sustaining growth. Yet little is known about them empirically, especially in developing nations. The purpose of this paper is to qualitatively and quantitatively characterize several types of the most commonly modeled externalities, and to examine their effects on welfare during trade liberalization.

1. *Types of Spillovers:*

Although there are many forms of external-economies, in this paper, I focus on industry-wide and localization externalities. I define industry-wide externalities as those externalities caused by a specific industry’s activity, independent of geography. By contrast, localization externalities are caused by a specific industry’s activity within a particular geographic area. These types have been in the literature the longest, are the best defined, and seem potentially the most important.

One of the hardest aspects of discussing spillovers, is distinguishing between their underlying causes. Ethier (1979) helps to clarify some of the differences between industry-wide and localization externalities. He points out that larger industries, independent of their firms' proximity, can experience externalities because of increased specialization—especially in a world with advanced transportation and communication systems. Larger industries can divide the production process into smaller steps that are performed in different plants, just as in Adam Smith’s famous example, pin production was broken down to very small steps within early factories.

While industry-wide externalities depend only on the size of the industry, localization externalities link productivity to both size and geographic concentration. According to Marshall (1890), the three main sources of industry-wide economies are: the development, attraction, and retention of specialized labor; the genesis of intermediate input producers; and more fluid exchanges of ideas and technology. These externalities figure prominently in both the endogenous growth and international trade literatures. Authors such as Rotemberg and Saloner (1990) and Krugman (1991a) have discussed reasons for workers and plants within industries to congregate together geographically. Specialized intermediate input producers are rigorously modeled by a number of authors such as: Ethier (1981), Helpman and Krugman (1985), Romer (1990), and Markusen (1990). Finally, knowledge spillovers, the last major force behind localization externalities, have been examined by many authors, including Griliches (1991), and Porter (1990).

2. *Empirical Evidence of External Economies:*

Despite their longtime presence in the writings of economists, little econometric work on externalities was done until interest was renewed by their prominent place in a series of endogenous growth models. One of the most important modern studies was done by Caballero and Lyons (1990). They used 3-digit European data to estimate economy-wide (caused by the total economic activity in the economy) externalities. They conclude that these types of external economies exist, and are substantial and positive. Bartlesman, Caballero and Lyons (1994) extended the study by estimating the external economies attributable to intermediate goods producers and customers with 4-digit manufacturing data. Their results indicate that in the short run industry-wide demand-based externalities are critical but in the long run intermediate goods producers are the primary source of external economies. Hanson (1994) uses 4-digit Mexican data to find that own-industry employment growth is positively related to the level of agglomeration of related industries. He also shows that own-industry agglomeration may negatively impact relative employment growth. Finally, he investigates the effects of the skill-mix of the local labor pool. His results indicate that industrial diversity has very little effect on employment growth. Finally, Jarmin (1997) uses plant-level U.S. manufacturing data to estimate a localization externality model that allows the degree of the spillover to vary with the geographic distance between the plant and the other plants in the industry.

With the exception of Jarmin’s (1997) work, most previous empirical studies of external economies have used aggregate data. Data at the two, or even four-digit level are not well suited to study external economies. Aggregate data do not allow researchers to disentangle the external and internal returns coefficients. Nor do they allow researchers to examine some of the more interesting types of externalities. Also, all of the studies that I am aware of use data from developed countries. The objective of this paper is to use plant-level longitudinal data from three semi-industrialized nations to provide some of the first micro evidence on the importance of several types of industry-wide and localization external economies in the developing world.

II. The Model:

My point of departure is the basic model developed in Caballero and Lyons (1990). My general estimation equations are[1]:

$$dy_{ijpt} = \gamma_j dx_{ijpt} + de_{jpt} + d\epsilon_{ijpt},$$

Here d's indicate first differences, lower-case letters indicate logs, y is value added, k is capital, l is labor, "$_j$ is the cost share of labor for industry j, e is an external economy index, v is an unobserved productivity index, and ,$_{1jt}$ is noise. Also:

$$dx_{ijpt} = \alpha_{jt} dl_{ijpt} + (1-\alpha_{cjt}) dk_{ijpt} ,$$

and

$$de_{jpt} = \beta_1 dz_{jpt} + \beta_2 dz_{jpt} G_j + d\epsilon_{ijpt} ,$$

where z_{jpt} is either labor or output[2] of the j^{th} industry in province p during year t, and G_j is a measure of industry agglomeration. Finally,

---

[1] I adopt the following notation: "I" denotes plants, "j" is for industry, "p" indexes province, and "t" is time.

[2] Potential misspecification because of the simultaneity between industry-wide output and plant-specific productivity shocks: corr(dy_{jpt},d:$_{ijpt}$) … 0, could bias the external returns to scale coefficients based on output. Using industry-level data, Caballero and Lyons (1990) show an analogous problem can be reduced by expressing aggregate output growth in terms of factor growth and productivity growth (of course y_{jpt}.and x_{jpt} are still simultaneously determined to the extent that the firms are affected by business cycles). In the same spirit, I substitute industry factor growth plus industry productivity growth for industry output growth.

$$\epsilon_{ijpt} = \mu_{ipj} + \tau_{jpt} + \xi_{ijpt} .$$

the error component $:_{ijp}$ is a plant-specific effect reflecting heterogeneous technologies and management; J_{jpt} is a time effect, common to all plants that reflects general changes in capacity utilization and technological innovation; and $>_{ijpt}$ is noise.

My work is distinct from theirs in three respects. First, I use plant-level data that allow me to examine external effects at the level that most theories predict they occur. Specifically, I estimate the effects of externalities from employment and output on individual plants. Second, I construct proxies for several types of externalities stressed by theory but not estimated by Caballero and Lyons: industry-wide, and localization. The final difference between my work and most other studies is that my data are from developing countries while most previous studies have featured developed countries.

IV. The Data

Three plant-level panel data sets from Chile, Mexico, and Morocco, spanning 7, 6, and 5 years respectively, are used to estimate the models. The Chilean data cover virtually all manufacturing plants with at least 10 workers observed at least once during 1979-1986. Outputs are deflated using price indices constructed from sectoral output prices using the 1977 Chilean input-output table. Capital stocks are imputed by applying the perpetual inventory method to deflated investment figures for each of four capital goods categories.[3] For more details, see Westbrook and Tybout (1993).

The Mexican data also comprise plant-level panels for several industries. They come from Mexico's Annual Industrial Survey and cover the period from 1984 through 1990. For an average industry, the data span approximately 80 percent of total output (the excluded plants are the smallest ones) and include information on: output, employment, location, input usage, costs, investment and inventories. Mexico's Secretary of Commerce and Industrial Development (SECOFI) provided industry-level deflators for output and intermediate inputs and sector-level deflators for machinery and equipment, buildings, and land.[4] A more detailed description of the data can be found in Tybout and Westbrook (1995).

The Moroccan data cover most manufacturing firms and span the years 1984-1989. Nominal variables are deflated using a set of sectoral price indices obtained from The World Bank. As with the Chilean data, capital stocks are imputed using the perpetual inventory method on deflated investment figures. The capital

[3]Base-year capital stocks are taken from 1980 financial statements and should reflect replacement costs.

[4]Maquiladora plants (plants that assemble components for export only) were excluded from the analysis because they do not report values for gross output or intermediate inputs.

stock for the base year, 1985, is established by multiplying sectoral capital/labor rates for firms with 10 or more employees by the number of employees. A perpetual inventory technique is used for the remaining years and a 5 percent depreciation rate of capital is assumed.

The data sets are too large to check the reliability of each observation. To eliminate outrageous values, the data are subject to a set of exclusion criteria. Valid observations require values greater than zero for: gross value of output, the capital stock, the number of employees, and the cost of labor. Additionally, observations with total costs (or gross value of output) per worker less than one twentieth or greater than twenty times the industry average are excluded. Also eliminated are observations showing either rates of growth of total cost (gross value of output) per worker greater than 300 percent per year or rates of decline of total cost (gross value of output) per worker greater than 75 percent per year.

Also, studentized residuals, the ratio of the residual to its standard error, are used to identify additional outliers. For each regression, observations that yield studentized residuals with absolute values greater than three are omitted and the regression is run again. The results remain qualitatively unchanged between the two stages in all of the plant-level regressions and the results reported here are from the second stage regressions.

Finally, all my estimations use "unbalanced" panels. Using balanced panels could bias the estimated increasing returns to scale (IRTS) upwards because new firms have higher failure rates than seasoned firms, and less-efficient firms fail more frequently. The IRTS coefficients estimated with balanced panels would be too high because the least-efficient plants are omitted. Using unbalanced panels mitigates this problem by increasing the heterogeneity of the pool of plants.

2. *Plant-Level Estimators:*

Recall from equation (4) that the error term of the production function, ϵ_{ijpt}, has three components that are unobservable to the econometrician:

$$\epsilon_{ijpt} = \mu_{ijp} + \tau_{jpt} + \xi_{ijpt} .$$

Here μ_{ijp} is a plant-specific effect, τ_{jpt} is a region and industry-specific time effect, and ξ_{ijpt}, is assumed to be identically independently distributed across plants and time and uncorrelated with the exogenous variables. The plant-specific effect, μ_{ijp}, can be removed with either a within or difference estimator. The within estimator is obtained by expressing the data in terms of deviations from plant-specific means and applying OLS to the transformed variables. That is, for any variable x, the within transformation is:

$$\hat{x}_i = x_{it} - (\frac{1}{T})\Sigma_{t=1}^{T} x_{it}, \quad i=1,...,n .$$

The jth-difference estimator results from applying OLS to variables transformed as follows:

$$d^j x_{it} = x_{it} - x_{it-j},$$

where d^j denotes the jth-difference operator. If there are T periods, any j value between 1 and T-1 may be chosen. An important distinction among the various estimators is sensitivity to measurement error (Griliches and Hausman (1986)). I report the results of first difference and within estimation, but I also comment on the effects of longer differences in a later section.

Both the within and difference estimators are based exclusively on the time variation within the data. To exploit cross-sectional variation, and to minimize measurement error bias, the between estimator is also employed. Although the between estimator has the advantage of focusing on cross-sectional variation, in this context it suffers some drawbacks. First, if the estimator is used on equations with industry-wide externality proxies, the estimated externality coefficients are likely to be biased. Since the variation exploited to compute these regressions is across industries and this level of externality proxy produces a single value for each industry, the externality proxies will also be picking up miscellaneous industry-specific effects. Second, the between estimator does not sweep out the plant-specific effects, $:_{ijp}$. This means that the estimated internal returns coefficients obtained with this estimator could be biased upward.

Since the individual externality proxies include variables that are correlated such as industry output and employment, and since only one proxy is used in each regression, the possibility of omitted variables bias exists. However, I choose to use only one proxy at a time for two reasons. First, because the proxies are often so closely correlated, some regressions containing multiple proxies fail due to near-perfect multicolinearity. Second, using one proxy at a time allows me to more closely follow the theoretical literature which usually specifies one type/level of externality in a particular model levels of externalities could be operating simultaneously. This allows me to discuss whether or not my findings support each model.

Finally, a common problem plaguing econometric work of this type is the obvious correlation of output and employment with demand: $corr(dy_{ijpt}, dJ_{jpt})$ … 0. Because of this, there is always a concern that the estimated "externalities" may actually be capturing capacity utilization effects. That is, because plants cannot costlessly adjust capital during business cycles, they often have excess capacity. Variables such as industry output that are correlated with demand, could appear to affect productivity by capturing these business cycle effects. Unfortunately, there is little that can be done to mitigate this problem. Although it is theoretically possible to control for time effects, J_{jpt}, by including year dummies in the models, because several of the externality proxies vary by year only, year dummies are not included, and the externality proxies can be expected to capture some of the time effects, J_{jpt}.

V. The Results

1. Industry-Wide Externalities

Many authors, including Pigou (1928) and Romer (1986)[5], use industry-wide externalities to motivate their trade and growth analyses. Tables 1- 6 report on the industry-wide external economy proxies, including industry output and employment obtained from the within and difference estimators. While the estimators performed similarly in the Combined and Chilean data, the within estimator produced a greater number of significant coefficients in the Mexican and Moroccan data. The distinction may be due to measurement error. I found that longer differences

[5]Although this model uses economy-wide, aggregate knowledge spillovers, I believe it captures the spirit of many of the own-industry models of externalities.

produced a greater number of significant coefficients in the Moroccan and Mexican data, and differences greater than one reduced the number of significant coefficients in the combined and Chilean data. It is not unreasonable therefore, to suspect that measurement error as the cause.

Industry-Wide Output:

Industry output should capture the combined effects of the three main sources of industry-wide externalities: more qualified labor, specialized intermediate inputs, and cross-plant knowledge spillovers. To estimate the combined force of these effects I specify the estimating equation as: $dy_{ijpt} = (_j\, dx_{ijpt} + \$_j dy_{jt} + ,_{ijpt}$. More sophisticated industries, such as Automobiles and Trucks or Pharmaceuticals and Medicines, should benefit most from these effects since they require more industry-specific skills and intermediate inputs. The results however, do not show a clear pattern among industries (Table 1). Yarn and Finishing of Fabric, a moderately sophisticated industry, as well as Automobile production, arguably the most sophisticated industry, display comparable evidence of external economies from industry output: almost all the significant coefficients from both estimators are, as modeled in the trade and growth literatures, positive. This evidence supports trade models like those developed by Graham (1923) Helpman (1984) and Helpman and Krugman (1985), general equilibrium models such as Markusen's (1990), and industrialization models like Lucas' (1988).

The equations estimated with the between estimator must be run across industries, producing a single coefficient for each data set. Table 2 reports the industry-wide output coefficient obtained from the between estimator. Generally, the coefficient is positive and significant, which aligns well with the results from the other two estimators. The coefficients obtained with the between estimator are noticeably larger than those obtained from the within or difference estimators. There are several possible explanations for this. One is that these coefficients are likely to be biased because of the externality proxy's correlation with other miscellaneous industry-specific effects. Another is that the between estimator exploits cross-sectional variation in the data, and it is quite plausible that this is the dimension over which the external economies are most apparent.

Table 1: Within and Difference Estimates of Industry-Wide Externalities from Output (6's in top row, SEs below; *= significant at approx 95%):

Industry	Within Estimates				Difference Estimates			
	Combined	Chile	Mexico	Morocco	Combined	Chile	Mexico	Morocco
Fruit & Veg	0.0136*	0.0561*	0.0439*	0.0211*	0.0376	0.0557*	0.1018	-0.0167
	0.0037	0.0134	0.0111	0.0082	0.0736	0.0273	0.0596	0.0125
Yarn, Fabric	-0.0026*	0.0074*	0.0027*	-0.0038*	0.0165*	0.0554*	0.0013	-0.0104*
	0.0006	0.0028	0.0012	0.0016	0.0027	0.0131	0.0026	0.0020
Taps & Carpet	0.0307	0.0770*	0.1211	0.0014	0.1605	0.2432*	0.1213	0.0768
	0.0276	0.0204	0.0965	0.0485	0.1933	0.0707	0.1294	0.0730
Non-Met Furniture	0.0104*	0.0020	0.0070*	0.0017	0.0281*	0.0060*	0.0103*	-0.0090
	0.0026	0.0018	0.0032	0.0174	0.0068	0.0029	0.0068	0.0086
Pharm & Meds	0.0024	-0.0220	0.0201*	-3.4833	0.0687*	-0.8435	0.0195*	-0.0069
	0.0041	0.0430	0.0058	35.4453	0.0150	0.9968	0.0078	0.0397
Soap, Perfume	-0.0105*	0.0011	0.0109*	0.0147	-0.0035	0.1486*	0.0061	0.0281
	0.0030	0.0147	0.0035	0.0976	0.0062	0.0686	0.0051	0.0256
Autos Trucks	0.0083*	0.0283*	0.0098*	0.0146	0.0123*	0.0904*	0.0115	0.0002
	0.0018	0.0052	0.0025	0.0099	0.0054	0.0270	0.0092	0.0120

Table 2: Between Estimates of Industry-Wide Externalities from Output (6's in top row, SEs below; *= significant at approx 95%):

Industry	Combined	Chile	Mexico	Morocco
All Inds	0.1641*	0.2930*	0.3157*	-0.2462
	0.0363	0.0314	0.0323	0.1714

Industry-Wide Employment:

Substituting total industry labor for industry output in the production functions ($dy_{ijpt} = (_{j}dx_{ijpt} + \$_{j}dl_{jt} + ,_{ijpt}$) helps focus on externalities from specialized labor. These externalities have been proposed in many trade models such as Ethier's (1979 and 1982) and Krugman's (1991a), and growth models such as Matsuyama's (1991). Again, more sophisticated industries are expected to benefit most from the availability of specialized labor. The results (shown in Tables 3 and 4) show that the evidence for externalities is at least as strong across all estimators, for industry labor as it is for industry output. Moreover, the Automobile and Furniture industries, among the more sophisticated gene industries studied rally have the largest number of significant coefficients across the data sets

Table 3: Within and Difference Estimates of Industry-Wide Externalities from Employment (6's in top row, SEs below; *= significant at approx 95%):

Industry	Within Estimates				Difference Estimates			
	Combined	Chile	Mexico	Morocco	Combined	Chile	Mexico	Morocco
Fruit & Veg	0.0143*	0.0118*	0.0334*	0.0142*	0.0058	0.0049	-0.0103	-0.0230*
	0.0023	0.0023	0.0061	0.0053	0.0046	0.0048	0.0325	0.0100
Yarn, Fabric	0.0035*	0.0067*	-0.0016	-0.0070*	0.0068*	0.0121*	-0.0055	-0.0061
	0.0008	0.0009	0.0041	0.0028	0.0014	0.0012	0.0055	0.0035
Taps & Carpet	0.0026	0.0622*	0.0999	0.0009	0.0045	0.1194*	-0.0247	0.0109
	0.0029	0.0131	0.1048	0.0030	0.0038	0.0216	0.1445	0.0128
Non-Met Furn	0.0130*	0.0101*	0.0167*	-0.0031	0.0102*	0.0087*	0.0213*	-0.0033
	0.0014	0.0014	0.0062	0.0111	0.0016	0.0017	0.0099	0.0087
Pharm & Meds	0.0392*	0.0058	0.0323*	0.0913*	0.0484*	0.0068	0.0478*	0.0159
	0.0056	0.0103	0.0069	0.0234	0.0070	0.0147	0.0087	0.0516
Soap, Perf	0.0148*	0.0046	0.0137*	-0.0044	-0.0019	0.0330*	0.0124	0.0213
	0.0037	0.0080	0.0042	0.0121	0.0085	0.0107	0.0124	0.0201
Autos Trucks	0.0134*	0.0150*	0.0144*	0.0083*	0.0137*	0.0308*	0.0079	0.0057
	0.0020	0.0028	0.0040	0.0039	0.0051	0.0041	0.0109	0.0102

Table 4: Between Estimates of Industry-Wide Externalities from Employment (6's in top row, SEs below; *= significant at approx 95%):

Industry	Combined	Chile	Mexico	Morocco
All Inds	0.3749*	0.5637*	0.4973*	.0.0587
	0.0589	0.0765	0.0763	0.1193

Industry-Wide Blue/White Collar Employment

Specifying the externality proxy as either blue or white collar industry-wide employment ($dy_{ijpt} = (_j\ dx_{ijpt} + \$_j dw_{jt} + 7_j db_{jt} + ,_{ijpt}\cdot_t$) helps identify the sources of industry employment effects. The results are displayed in Tables 5 and 6. Notably, the blue collar employment coefficients from all three estimators are more likely to be significant and are almost always positive while the white collar coefficients are often negative. The coefficients obtained with the between estimator also show this pattern. These results support the blue/white collar externality distinction drawn by Hanson (1992). They suggest that the industries' blue collar workers possess many of the specialized skills that create positive externalities.

2. Localization Economies:

I use two methods to test for localization economies. First, I construct a measure of overall industry concentration and interact it with output or employment. Second, I measure

Table 5: Within and Difference Estimates of Industry-Wide Externalities from Blue/White Collar Emp (6's in top row, SEs below; *= significant at 95%):

Ind	Within Estimates						Difference Estimates					
	Combined		Chile		Mexico		Combined		Chile		Mexico	
	Blue	White	Blue	White	Blue	White	Blue	White	Blue	White	Blue	White
Fruit	.0153*	-.0048	.0024	.0100*	.0248*	.0159	-.0061	.0093*	-.0015	.0087	-.0048	.0157
Veg	00038	.0038	.0052	.0048	.0068	.0236	.0052	.0047	.0059	.0051	.0246	.0344
Yarn	.0061*	-.0044	.0040*	.0041*	-.0033	-.0080	.0079	-.0041	.0148*	-.0077	-.0022	-.001
Fab	.0009	.0010	.0011	.0013	.0032	.0020	.0014	.0016	.0014	.0021	.0062	.0025
Taps	.0683*	-.0017	.0643*	.0156	.0620	.0057	.0998*	.0090	.1150*	-.0096	-.0129	-.023
Carp	.0142	.0304	.0139	.0307	.0893	.0770	.0208	.0147	.0236	.0143	.1332	.0937
Furn	.0095*	.0030*	.0112*	-.0040	-.0353*	.0279*	.0065*	-.0014	.0077*	-.0027	.0103	.0097
	.0020	.0019	.0020	.0020	.0150	.0078	.0018	.0014	.0019	.0014	.0146	.0099
Phar Med	.0327*	.0118*	.0208*	.0127*	.0411*	-.0306	.0293*	.0165*	.0032	.0140	.0313*	-.017
	.0039	.0028	.0107	.0049	.0061	.0106	.0039	.0048	.0119	.0072	.0051	.0135
Soap Perf	.0179*	.0051	.0065	-.0008	.0108	.0021	.0079	.0073	.0164*	.0212*	.0105	.0002
	.0057	.0035	.0069	.0058	.0092	.0053	.0070	.0061	.0082	.0071	.0110	.0098
Auto	.0152*	.0021	.0232*	.0234*	.0230*	-.0085	.0092*	.0229*	.0268*	.0226*	.0069	-.006
	.0022	.0029	.0046	.0095	.0116	.0110	.0045	.0080	.0036	.0061	.0922	.0169

Table 6: Between Estimates of Industry-Wide Externalities from Blue/White Collar Emp (6's in top row, SEs below; *= significant at approx 95%):

Industry	Combined		Chile		Mexico	
	Blue	White	Blue	White	Blue	White
All Inds	-0.2368*	0.3159*	0.3243*	-0.3990*	0.4229*	-0.1480*
	0.0645	0.0246	0.0408	0.1371	0.0366	0.0658

industry output or employment within a plant’s province and use that as my externality proxy. To construct a country-specific measure of industry agglomeration, I use a geographic GINI index developed in Krugman (1991a). The index is created by measuring the area between a 45 degree line and a curve made by plotting cumulative manufacturing employment against cumulative industry employment by province. The index varies between zero (least concentrated) and one-half (most concentrated). The geographic GINI coefficients are reported in Table 7 (also found in the appendix) by country and industry.

The GINIs are slightly smaller than those in Krugman (1991a) but provide some evidence that many individual industries are geographically concentrated. Most of the GINIs are well above zero. Furthermore, several industries, such as Tapestries and Carpets, and Soap, Perfumes and Toiletries, have moderately high GINIs in all three countries. Hereafter, when using plant-level data, I will focus on the following industries: Fruit & Vegetable Canning; Yarn, Finishing of Fabric; Tapestries & Carpets; Non-metal Furniture; Pharmaceutical & Medicines; Soap, Perfumes & Toiletries; and Automobiles & Trucks. These industries are chosen because of their moderate to large GINI coefficients and anecdotal evidence of their agglomeration.

Table 7: Location GINI Coefficients:

Ind #	Industry	Chile		Mexico		Morocco	
		# Plants	GINI	# Plants	GINI	# Plants	GINI
1	Slaughter, Preparation of Meats	42	0.11	46	0.21	3	0.38
2	Dairy Products	17	0.23	26	0.22	13	0.30
3	Fruit and Vegetable Canning	25	0.18	23	0.31	43	0.28
4	Preparation and Preservation Seafood	19	0.26	20	0.42	33	0.34
5	Animal and Vegetable Products	19	0.25	36	0.24	46	0.32
6	Milled Grains	42	0.14	88	0.21	85	0.22
7	Bakery Goods	502	0.06	22	0.26	390	0.18
8	Cocoa, Chocolates and Confections	12	0.22	9	0.38	14	0.21
9	Animal Feeds	5	0.12	32	0.21	33	0.28
10	Distillation of Alcoholic Beverages	9	0.26	12	0.34	0	0.00
11	Wine and Brandy	28	0.15	10	0.30	9	0.40
12	Beer and Malt	2	0.15	16	0.27	1	0.43
13	Non-Alcoholic Beverages and Soda	13	0.08	68	0.20	11	0.29
14	Tobacco Products	1	0.30	7	0.32	1	0.44
15	Yarn, Fabric and Finishing of Textiles	67	0.15	112	0.19	72	0.17
16	Articles Made of Textiles but not Clothes	11	0.18	13	0.29	39	0.22
17	Tapestries and Carpets	6	0.31	5	0.41	29	0.34
18	Fabrication of Clothes Except Shoes	130	0.23	96	0.24	195	0.19
19	Shoe Manufacturing	65	0.17	44	0.37	68	0.25
20	Non-Metal Furniture	45	0.14	39	0.25	15	0.29
21	Wood Pulp, Paper and Cardboard	6	0.22	42	0.24	5	0.40
22	Paper and Cardboard Boxes and Containers	5	0.24	16	0.25	37	0.21
23	Printing and Publishing	87	0.13	62	0.35	159	0.20
24	Basic Industrial Chemicals not Fertilizer	14	0.17	58	0.23	3	0.38
25	Fertilizers and Pesticides	1	0.26	22	0.24	9	0.40
26	Syn Resins, Plastics and Art Fibers not Glass	2	0.18	31	0.24	2	0.37
27	Paints, Varnishes and Lacquers	16	0.22	43	0.21	14	0.31
28	Pharmaceuticals and Medicines	29	0.27	69	0.34	15	0.30
29	Soap, Perfumes and Toiletries	20	0.28	40	0.32	22	0.26
30	Tires	9	0.18	9	0.27	6	0.37
31	Non-Tire Rubber Products	19	0.25	36	0.27	6	0.37
32	Plastic Products	65	0.24	68	0.22	79	0.23
33	Ceramics, Pottery, and Clay Const Mater	6	0.21	16	0.35	42	0.28
34	Glass and Glass Products	12	0.22	18	0.29	10	0.38
35	Cement, Lyme, Gypsum and Plaster	4	0.20	88	0.20	13	0.32
36	Non-Metal Mineral Products	4	0.21	10	0.28	76	0.26
37	Iron and Steel	10	0.23	50	0.28	2	0.34
38	Lead, Zinc, Tin, and Nickel	4	0.27	12	0.22	4	0.31
39	Hand Tools and Cutlery	12	0.24	8	0.40	13	0.28
40	Metallic Furniture Except Electric Lamps etc	9	0.22	32	0.31	4	0.29
41	Structural Metal Products	26	0.13	36	0.22	59	0.20
42	Agricultural Machines and Equipment	7	0.20	7	0.40	2	0.41
43	Spec Indus Mach not for Wood/Metal Working	1	0.28	31	0.23	3	0.31
44	Office Machines, Adding Machines and Equip	3	0.31	3	0.32	1	0.31
45	Industrial Electrical Machines and Equip	4	0.17	40	0.25	11	0.23
46	Radios, Television and Common Equip	1	0.31	28	0.24	12	0.27
47	Domestic Electrical Machines and Equip	5	0.26	16	0.23	0	0.38
48	Shipbuilding and Repair	2	0.48	0	0.00	10	0.42
49	Railroad Equipment	10	0.29	7	0.42	1	0.42
50	Automobiles and Trucks	26	0.17	31	0.22	25	0.26
51	Motorcycles and Bicycles	2	0.27	10	0.28	6	0.30
52	Photographic and Optical Equipment	4	0.23	3	0.33	1	0.41

A) The Effects of the GINIs:

When I include a geographic GINI index, interacted with the externality term in the estimation equation, the estimating equation takes the following general form: dy_{ijpt} = $(_j$ dx_{ijpt} + $\$_j$ (GINI*de_{jt}) + $8_j de_{jt}$ + $,_{ijpt}$. The most striking feature of the GINIs' coefficients is that, as in Hanson (1994), they

Table 8: Within and Difference Estimates of Coefficients from GINI Interactions (6's in top row, SEs below; *= significant at approx 95%):

Industry	Within Estimates				Difference Estimates			
	G*Out	G*Emp	G*Bl	G*Wh	G*Out	G*Emp	G*Bl	G*Wh
Fruit & Veg	-0.1192*	0.1365*	0.3256*	-0.1501	-0.1240	-0.0333	0.0972	0.2690
	0.0594	0.0420	0.0656	0.1864	0.0811	0.0965	0.1955	0.2894
Yarn, Fabric	-0.1709*	-0.3709*	-0.1714*	-0.4896*	-0.5456*	-0.6809*	-0.5973*	0.1666*
	0.0340	0.0826	0.0834	0.0611	0.1097	0.1263	0.1452	0.0837
Taps & Carpet	-3.4122*	-2.6055*	0.1646	-1.4336	-3.7852*	-2.6629*	-1.1862	0.5927
	0.5167	0.4582	1.1321	1.0255	0.6854	0.5378	1.1939	0.7760
Non-Met Furn	-0.0661	0.1780*	-0.0224	0.2653*	-0.0971*	-0.0343	-0.0315	0.1640*
	0.0400	0.0426	0.1463	0.0763	0.0436	0.0524	0.1323	0.0911
Pharm & Meds	-0.2722*	0.5341*	0.4032*	-0.8617*	0.8634*	0.5020	0.1870	-0.5638*
	0.1507	0.1821	0.1884	0.1583	0.2226	0.2794	0.2259	0.1883
Soap, Perf	0.2018	0.8530*	0.5442	0.4987*	1.0091*	0.8237*	0.5827	-0.3713
	0.1649	0.1683	0.3120	0.2237	0.2452	0.3042	0.3841	0.3394

Autos Trucks	-0.2768*	-0.1011*	0.0417	-0.6481*	-0.5286*	-0.3660*	-0.4673*	-0.4905
	0.0839	0.0563	0.2294	0.2899	0.1706	0.1645	0.1891	0.3129

are frequently negative. Ellis and Fellner (1943) suggest two potential causes of negative externalities. First, diminishing returns can be due to the presence of an industry-specific factor with a fixed supply. Second, increasing transfer costs of a factor that is used by multiple industries at ever increasing prices can adversely affect productivity. Of course, a combination of the two causes could occur. Other authors (David and Rosenbloom (1990)) have noted that external diseconomies can be caused by "congestion costs". That is, as a region becomes more crowded, the cost of adding additional units of capital increases and the benefit of additional units of local labor decreases. Eventually the marginal congestion costs equal the positive externalities and new plants no longer enter the region.

B) Own-Province Industry Activity

While the GINI interactions help capture cross-country variation in industry concentration, they do not use all of the available information in the plant-level data. By measuring the output of the plant's industry within its province, I come closer to measuring the effects of industry activity in the plant's immediate vicinity. The GINIs measure relative agglomeration while province-specific, industry-wide activity directly measures the volume of local, own-industry production. Another advantage of using this externality specification ($dy_{ijpt} = (_j\, dx_{ijpt} + \$_j de_{jpt} + ,_{ijpt}$) is that it allows me to more fully employ the between estimator. Because the localization externality proxies vary by province and industry, regressions using this estimator can be run within individual industries just as they are for the other two estimators.

Localization Externalities from Output:

Many of the province-specific, industry-wide output coefficients (see Table 9 for within and difference estimates, Table 10 for between estimates.) obtained with all three estimators are significant, and positive. This supports the many traditional models of external economies of scale in trade, growth, and urban economics that Dierx (1990) surveys.

Table 9: Within and Difference Estimates of Localization Externalities from Output (6's in top row, SEs below; *= significant at approx 95%):

Industry	Within Estimates				Difference Estimates			
	Combined	Chile	Mexico	Morocco	Combined	Chile	Mexico	Morocco
Fruit	0.0122*	0.0070	0.0215	0.0112	-0.3947	-0.0347	-0.0125	0.0044
Veg	0.0061	0.0128	0.0211	0.0100	0.5105	0.0278	0.0647	0.0526
Yarn	0.0009	0.0068*	0.0076*	-0.0109*	0.01055	0.0385*	0.0002	-0.0077
Fabric	0.0014	0.0028	0.0032	0.0035	0.0263	0.0091	0.0049	0.0054
Taps	-0.0780*	0.0547*	0.0091	-0.0157	0.0055	0.1104*	-0.0321	-0.0839
Carpet	0.0116	0.0139	0.0307	0.0174	0.2243	0.0349	0.0459	0.1112
Non-Met	0.0189*	0.0101*	0.0143*	0.0453*	0.0490*	0.0150*	0.0162	-0.0069
Furn	0.0033	0.0037	0.0069	0.0279	0.0107	0.0055	0.0117	0.0219
Pharm &	0.0115*	-0.0611	0.0196*	-0.0199	0.0508*	-0.2202	0.0127	0.0187
Meds	0.0038	0.0496	0.0063	0.0121	0.0137	0.2456	0.0078	0.0473
Soap,	0.0059	-0.0056	0.0163*	0.0119	0.0072	0.0832*	0.0019	0.0047
Perf	0.0048	0.0116	0.0049	0.0158	0.0045	0.0365	0.0042	0.0562
Autos	0.0198*	0.0202*	0.0129*	0.0236	0.0094*	0.0585*	0.0052	0.0143
Trucks	0.0036	0.0059	0.0048	0.0124	0.0035	0.0198	0.0039	0.0169

Note the contrast between these results and the findings on concentration from the GINI index (and could be proxying capacity utilization effects). Some industries (Carpets, Automobiles), which show negative effects from the GINIs, show positive effects from increased local output. This may be because while the GINIs measure cross-country agglomeration (which would likely be easily affected by forces that cause negative externalities), this metric quantifies the effects of increased local production, and local production does not give any information about the overall concentration of the industry.

Table 10: Between Estimates of Localization Externalities from Output (6's in top row, SEs below; *= significant at approx 95%):

Ind	Combined	Chile	Mexico	Morocco
Fruit & Veg	-0.5036	-0.6355	-0.2834	-0.3527
	0.3962	0.4127	0.4080	0.5459
Yarn, Fabric	-0.1423	-0.4565	0.0109	-0.0789
	0.0877	0.2839	0.0893	0.1256
Taps Carpet	-0.8231	0.9940*	0.0395	-1.4457
	0.8924	0.0023	0.0749	2.3073
Non-Met Furn	0.0964	-0.0965	0.4025*	0.1100
	0.0636	0.1670	0.0972	0.1128
Pharm & Meds	0.1063	1.1967*	0.1637	-0.5197
	0.1263	0.1333	0.1316	0.6361
Soap, Perfume	0.0528	0.5081*	0.0638	-0.0951
	0.1292	0.0752	0.1159	0.3979
Autos Trucks	0.0594	0.0479	0.1158	0.0302
	0.0583	0.0743	0.0584	0.1599

Localization Externalities from Employment:

Localization externalities from employment are widely used in trade and growth models such as those by Krugman (1991a), (1991b). To test for the existence of these externalities, I specify the estimation equation as: $dy_{ijpt} = (_j\ dx_{ijpt} + \$_j dl_{jpt} + ,_{ijpt}$. Tables 11 and 12 show that the evidence for. employment localization externalities is strong. Most of the coefficients are significant, and positive.

Table 11: Within and Difference Estimates of Localization Externalities from Employment (6's in top row, SEs below; *= significant at approx 95%):

Industry	Within Estimates				Difference Estimate			
	Combined	Chile	Mexico	Morocco	Combined	Chile	Mexico	Morocco
Fruit Veg	-0.0026	0.0101	0.0189	-0.0082	0.00093	-0.0098	-0.0025	0.0036
	0.0072	0.0061	0.0310	0.0142	0.009737	0.0110	0.0466	0.0191
Yarn Fabric	0.0039*	0.0070*	0.0036	0.0006	0.0069*	0.0123*	-0.0063	-0.0016
	0.0011	0.0010	0.0051	0.0055	0.0015	0.0013	0.0058	0.0056
Taps Carpet	-0.0090	0.0731*	0.0049	-0.0381	0.0377*	0.1124*	0.1360	-0.0357
	0.0148	0.0143	0.1680	0.0231	0.0155	0.0213	0.1964	0.0302
Non-Met Furnit	0.0166	0.0171*	0.0203*	0.0273	0.0209*	0.0174*	0.0250	-0.0196
	0.0238	0.0023	0.0098	0.0245	0.0033	0.0032	0.0145	0.0234
Pharm Med	0.0124	-0.0183	0.0280*	-0.0715*	0.0362*	0.0021	0.0350*	-0.0121
	0.0083	0.0131	0.0104	0.0322	0.0082	0.0141	0.0100	0.0525
Soap, Perf	-0.0014	-0.0020	0.0317*	-0.0523	0.0031	0.0290*	0.0005	-0.1417*
	0.0078	0.0093	0.0093	0.0331	0.0074	0.0123	0.0076	0.0522
Autos Trucks	0.0229*	0.0219*	0.0170*	0.0250*	0.0102*	0.0391*	0.0062	0.0026

	0.0039	0.0058	0.0062	0.0107	0.0037	0.0082	0.0044	0.0189

It is interesting to note that Fruit and Vegetable Canning does not show much evidence of either output or employment based localization externalities except in the between estimations where it is usually negative. This may help explain why there are fewer examples of geographic concentration in this industry than the other industries I selected for this study.

Table 12: Between Estimates of Localization Externalities from Employment (6's in top row, SEs below; *= significant at approx 95%):

Industry	Combine	Chile	Mexico	Morocco
Fruit & Veg	0.1977	-0.4430*	-0.4451	0.3442*
	0.1457	0.2192	0.3788	0.1771
Yarn, Fabric	-0.0048	-0.2800	0.0354	0.0689
	0.0711	0.1495	0.0926	0.1108
Taps & Carpet	0.1882	-17.1470*	0.0452	0.3294
	0.2599	6.3461	0.0879	0.3847
Furnitur e	0.1814*	-0.1940	0.5507*	0.3199*
	0.0828	0.1932	0.1843	0.1466
Pharm & Meds	0.0744	-3.0860	0.2528	0.0356
	0.1903	1.7434	0.2606	0.3104
Soap, Perf	0.2191	1.3300*	0.1645	0.2226
	0.1440	0.3792	0.1388	0.2955

Autos	0.1391*	0.0840	0.1523	0.2253
Trucks	0.0864	0.1356	0.0908	0.1999

Localization Externalities from Blue/White Collar Employment:

Externalities from the employment levels of different classes of workers is discussed in Hanson (1992). He shows that firms in an industry (textiles) may distinguish between some types of workers (white collar) who provide externalities when locally abundant, and others who may have more generic, easily learned skills (blue collar). In this section I examine this distinction by letting localized industry blue or white collar employment serve as the proxy for external returns to scale ($dy_{ijpt} = (_j\ dx_{ijpt} + \$_j dw_{jpt} + 7_j db_{jpt} + ,_{ijpt}$).

Table 13: Within and Difference Estimates of Localization Externalities from Blue/White Employment (6's in top row, SEs below; *= significant at 95%):

Ind	Within Estimates						Difference Estimates					
	Combined		Chile		Mexico		Combined		Chile		Mexico	
	Blue	White	Blue	White	Blue	White	Blue	White	Blue	White	Blue	White
Fruit Veg	.0306*	-.0272*	.0390*	-.0227*	.0142	.0215	.0085	-.0207*	.0095	-.0183*	.0185	-.0998
	.0111	.0079	.0120	.0084	.0360	.0891	.0125	.0087	.0135	.0093	.0439	.0943
Yarn Fabric	.0016	.0032*	.0038*	.0051*	.0069	-.0056	.0091*	-.0055*	.0151*	-.0081*	-.0054	-.0010
	.0011	.0013	.0012	.0014	.0049	.0035	.0017	.0023	.0016	.0022	.0056	.0047
Tap Carpet	.0332	-.0204	.0114	-.0423*	-.0589	-.1792	.1088*	.0145	.1155*	-.0166	.0435	.1608
	.0282	.0192	.0321	.0214	.1417	.1237	.0229	.0133	.0241	.0140	.1663	.1507
Non-Met Fur	.0027	.0095*	.0140*	.0003	-.0388	.0408*	.0078	.0030	.0108*	-.0002	.0166	.0097
	.0042	.0037	.0045	.0042	.0213	.0123	.0041	.0028	.0043	.0030	.0198	.0129
Pharm Meds	.0228*	.0118*	-.0076	-.0040	.0432*	-.0332*	.0237*	.0126*	-.0029	.0101	.0301	-.0148
	.0052	.0046	.0113	.0087	.0103	.0126	.0055	.0064	.0125	.0070	.0092	.0156
Soap, Perf	.0066	.0019	.0027	-.0069	.0106	.0145	-.0091	.0152	.0125	.0242*	-.0088	.0124
	.0059	.0049	.0082	.0061	.0083	.0075	.0063	.0079	.0090	.0073	.0102	.0143
Autos Trucks	.0179*	-.0021	.0132	-.0074	.0191	-.0059	.0037	.0042	.0353*	.0182*	-.0012	.0072
	.0037	.0043	.0069	.0081	.0145	.0164	.0056	.0066	.0073	.0070	.0131	.0158

Table 14: Between Estimates of Localization Externalities from Blue/White Collar Employment (β's in top row, SEs below; *= significant at approx 95%):

Industry	Combined		Chile		Mexico	
	Blue	White	Blue	White	Blue	White
Fruit & Veg	-0.6151*	0.2260	-0.5320	0.1310	-0.7156	0.2948
	0.2057	0.1466	0.2601	0.2021	0.3626	0.2258
Yarn Fabric	-0.2251	0.1520	-0.3210	0.0270	-0.1950	0.2550*
	0.1458	0.1221	0.6168	0.4870	0.1355	0.1174
Taps & Carpet	-0.4410*	0.3980*	0.0000	0.4450*	0.0691	-0.0191
	0.2142	0.1328	0.0000	0.1646	0.5265	0.4393
Furniture	0.0906	-0.0255	-0.1300	-0.0410	0.4825	0.0675
	0.1510	0.0569	0.1944	0.0597	0.3202	0.2321
Pharm & Meds	-0.1151	0.5111*	0.0000	0.4580	-0.4626	0.7267*
	0.2478	0.1523	0.0000	0.2589	0.3937	0.2523
Soap, Perf	These	-0.3124	1.2790	0.1970	0.6420	-0.3268
	0.2898	0.2344	0.5427	0.4563	0.3403	0.2747
Autos Trucks	-0.1433	0.2401*	-0.3370	0.2820	-0.0489	0.2076
	0.1441	0.1187	0.3217	0.1969	0.1773	0.1780

My results accord well with the belief that there is a distinction between the two types of workers, and suggest that the difference may actually be fairly strong. The evidence shows that in some industries an abundance of white collar workers may hinder a firm's productivity, but the presence of skilled blue collar workers is often helpful. (Of course blue-collar jobs may be more pro-cyclical than white-collar jobs which means that this pattern would also appear if my externality proxies are picking up capacity utilization effects due to business cycles.)

VI. Conclusions:

The evidence on external economies of scale reported here supports the theoretical literature, aligns well with previous empirical studies. However, all such studies are vulnerable to the possibility of merely reflecting capacity utilization effects. I found evidence of industry-level output and employment having a positive impact on plant productivity. Concentrations of own-industry workers, especially blue collar workers, enhances productivity.

Geography appeared to be an important consideration, though my results were mixed. Own-industry activity within a plant's province showed evidence of positively impacting productivity. However, industry concentration, as measured by geographic GINI coefficients, produced several negative coefficients. Taken together, these findings could indicate that a high volume of local output enhance productivity while high levels of industry agglomeration may hinder it (possibly because of congestion costs or the rationing of a scarce input). Given the simultaneity between city size and externalities, this is a plausible finding. Local industry output is more likely to pick up externalities than it is congestion costs while the GINI's are designed to measure agglomeration which is an excellent proxy for congestion. The negative coefficient on the GINI*industry output coefficient could be signaling that producing in congested areas is costly.

Another explanation for the sign difference is that the own-industry output coefficients, which are more likely to reflect capacity utilization, are biased. If own-industry output were capturing capacity utilization effects, the coefficients would be positive, while the externalities could cause the coefficients to be either positive or negative. Since the GINIs are less likely to proxy capacity utilization, and are often negative, it is possible that the sign difference indicates that the own-industry output coefficients are biased.

BIBLIOGRAPHY

Bartlesman, E., Caballero, R. and R. Lyons (1994) "Customer and Supplier-Driven Externalities", *American Economic Review* Vol 84, No. 4.

Caballero, R., and R. Lyons (1990) "Internal Versus External Returns to Scale in European Industry," *European Economic Review*, pp. 805-830.

Ethier, W. (1979) "Internationally Decreasing Costs and World Trade," *Journal of International Economics* No. 9 pp. 1-24.

Ellis, H.S., and W. Fellner (1943) "External Economies and Diseconomies," *American Economic Review*, Vol 33 pp. 493-511: reprinted in G.J. Stigler and K.E. Boulding, eds., *Readings in Price Theory*, pp. 242-263.

Grilliches, Z., and J.A. Hausman (1984) "Errors-in-Variables in Panel Data," *NBER Technical Paper No. 37*.

Hanson, G. (1992) "Agglomeration, Dispersion, and the Pioneer Firm: Theory and Evidence from Mexico," Mimeo.

Hanson, G. (1994) "Regional Adjustment to Trade Liberalization," Mimeo.

Hausman, J.A. (1978) "Specification Tests in Econometrics," *Econometrica* Vol 46, pp. 1251-1271.

Helpman, E. (1984) "Increasing Returns, Imperfect Markets and Trade Theory." in R. Jones, ed., *The Handbook of International Economics* North-Holland, New York, NY.

Helpman, E. and P. Krugman (1985) *Market Structure and Foreign Trade: Increasing Returns, Imperfect Competition, and the International Economy* MIT Press, Cambridge, MA.

Jarmin, R. (1997) *Learning by Doing and Plant Location,* Mimeo.

Krugman, P. (1991a) *Geography and Trade* MIT Press, Cambridge.

Krugman, P. (1991b) "History Versus Expectations," *Quarterly Journal of Economics* pp. 651-667.

Lucas, R.E. (1988) "On the Mechanics of Economic Development", *Journal of Monetary Economics* Vol 22, pp. 3-42.

Markusen, J.R. (1990) "Micro-foundations of External Economies", *Canadian Journal of Economics* Vol. 23, pp. 495-508.

Marshall, A. (1890) eighth edition, *Principles of Economics: An Introductory Volume.* MacMillan Company, New York.

Matsuyama, K. (1991) "Increasing Returns, Industrialization, and Indeterminacy of Equilibrium," *Quarterly Journal of Economics* pp. 617-650.

Pigou, A. (1928) "An Analysis of Supply," *The Economic Journal* No. 150 pp. 238-257.

Porter, M. (1990) *The Competitive Advantage of Nations* The Free Press, N.Y.

Romer, P. (1986) "Increasing Returns and Long-Run Growth", *Journal of Political Economy* Vol 94, pp. 1002-1037.

Rotemberg, J., and G. Saloner, (1990) "Competition and Human Capital Accumulation: A Theory of Interregional Specialization and Trade," *NBER Working Paper No. 3228*.

Tybout, J.R., and M.D. Westbrook (1995) "Trade Liberalization and Dimensions of Efficiency Change in Mexican Manufacturing Industries," *Journal of International Economics*.

Westbrook, M.D., and J.R. Tybout, (1993) "Estimating Returns to Scale with Large Imperfect Panels: An Application to Chilean Manufacturing Industries," *World Bank Economic Review* January.